GB04/04907

The Patent Office

PATENTS • DESIGNS • COPYRIGHT • TRADE MARKS

REC'D 20 DEC 2004

WIPO PCT

INVESTOR IN PEOPLE

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed Andrew Gersey

Dated 30 November 2004

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

24NOV03 E854166-17 D00056
P01/7700 0.00-0327179.8

1/77

The Patent Office

Cardiff Road
Newport
South Wales
NP9 1RH

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

THE PATENT OFFICE
21 NOV 2003
RECEIVED BY HAND

1.	Your reference	P35460GB/PWC
2.	Patent application number *(The Patent Office will fill in this part)*	0327179.8
3.	Full name, address and postcode of the or of each applicant *(underline all surnames)*	UWS Ventures Limited The Innovation Centre University of Wales Swansea Singleton park, Swansea, SA2 8PP
	Patents ADP number *(if you know it)*	8036394001
	If the applicant is a corporate body, give the country/state of its incorporation	United Kingdom
4.	Title of the invention	Purification method
5.	Name of your agent *(if you have one)* "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	Kilburn & Strode 20 Red Lion Street London WC1R 4PJ
	Patents ADP number *(if you know it)*	125001
6.	If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number	Country — Priority application number *(if you know it)* — Date of filing *(day / month / year)*
7.	If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application — Date of filing *(day / month / year)*
8.	Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:* *a) any applicant named in part 3 is not an inventor, or* *b) there is an inventor who is not named as an applicant, or* *c) any named applicant is a corporate body.* *See note (d))*	YES

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	28
Claim(s)	2
Abstract	8
Drawing(s)	5 ×5

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature Date 21 November 2003

12. Name and daytime telephone number of person to contact in the United Kingdom

Paul Chapman
Tel: 020 7539 4200

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

Purification method

The present invention relates to a method for purifying a recombinant glucose binding protein, in particular the lectin Concanavalin A (Con A). The method specifically utilises a buffer in which impurities, such as glycogen and other substances are soluble, but in which the protein remains insoluble. The use of such buffers, and the purified proteins are also described.

Con A is the best-characterised member of a large family of homologous carbohydrate-recognising proteins abundant in the seeds of leguminous plants [1]. Con A was the first of the legume lectins to be isolated [2] and is specific primarily for mannose, glucose and polysaccharides carrying these as α-linked non-reducing terminal sugars [1]. The saccharide specificity of Con A (found in the Jackbean *Canavalia ensiformis*) is similar to that of the closely related lectins from pea (*Pisum sativum*), lentil *(Lens culinaris)* and broad bean *(Vicia faba)*. In addition to these members of the "Glucose/Mannose Group", there are many other lectins from leguminous and other sources with very different specificities [1].

The natural plant source is rich in Con A (about 10% by weight of the Jackbean seed) and a saline extraction process produces crystallisable lectin [2]. However, the well-established method of purification is by affinity chromatography utilising the bio-specificity of this protein. Con A was the first lectin to be isolated by this technique. It binds readily to a cross-linked dextran gel (primarily an α-(1→6)-linked glucose polymer in beaded form, e.g. Sephadex G-75™ and may be

specifically eluted with a cognate monosaccharide for example, glucose, mannose or derivatives [3, 4]. This straightforward affinity purification methodology is widely used to produce commercially available native Con A and related lectins from their natural sources (*e.g. Sigma-Aldrich Co.*, Poole, UK: Biochemicals and Reagents Catalogue 2002-03).

At neutral pH, Con A is a tetramer of non-glycosylated subunits each of a molecular weight of 25,600 Da, but preparations of native Con A contain variable amounts of two smaller polypeptide fragments in addition to these intact subunits [1]. These fragments are the result of incomplete post-translational processing involving proteolytic cleavages, re-ligation and circular sequence permutation of precursor proteins during the complex biosynthesis of this lectin in the developing seed (summarised in [5] and [6]). The mature form of Con A is circularly permutated with respect to legume lectins. However, the amino acid sequence derived from the cDNA for the Con A precursor (pro-Con A) has direct, and not circular homology with other legume lectins. Therefore, a post-translational transposition and ligation (at mature protein residues 118 and 119) of two polypeptides must occur. Furthermore, the pro-Con A made in the plant is glycosylated, and unable to bind to cross-linked dextran. The oligosaccharide is removed prior to the proteolytic cleavages during post-translational processing in the plant.

It is desirable to express the various forms of this protein from cloned DNA sequences. This would allow the processing intermediates which are difficult to isolate from the plant to be obtained, in order to study the mechanism of their

biosynthesis. It would also enable the structure and/or function of any form to be manipulated by protein engineering (site-directed mutagenesis). In addition a secure and (in principle) unlimited and highly consistent supply of any form of the protein would be produced for any purpose for which this might be required (*e.g.* FDA requirements).

For many current uses the natural and commercially available Jackbean-derived lectin may be adequate. However, recombinant mature Con A, expressed directly in bacteria from re-ordered segments of precursor DNA, has not undergone the plant biosynthetic process and hence is uncontaminated with the smaller polypeptide fragments mentioned earlier. This recombinant protein might thus be superior to the native protein for some applications.

The first forms of recombinant Con A expressed corresponded to precursors present in developing seeds: pre-pro-Con A [7] and pro-Con A (non-glycosylated) [5]. Later forms included subunit fragments (A-chain and B-chain) [8], mature lectin [6] and various further derivatives of these proteins. Active recombinant forms of Con A following expression in bacteria (*Escherichia coli*) have been purified utilising the process described by Min *et al.* [5]. This procedure is summarised below from the point of harvesting of bacterial cells that had earlier been induced to express and accumulate recombinant lectin in liquid culture.

Cells were collected and resuspended in 50 ml MOPS-metals buffer (20 mM MOPS [3-(*N*-morpholino)propanesulphonic acid] pH 7.0, 1 mM $CaCl_2$, 1 mM $MnCl_2$, 1 M NaCl) and lysed by sonication (3 x 30 s) on ice. The lysate was placed on ice for 1 h followed by centrifugation at 27 000 X *g* for 30 min at 4 °C. The supernatant was

used as a soluble extract and subjected directly to dextran affinity chromatography (see below). However, the major portion of recombinant lectin was found in the pelleted debris after cell lysis and this insoluble material was refolded as follows:

The pelleted cell debris was resuspended in 2 ml denaturant (7 M guanidine hydrochloride in MOPS-metals buffer) and placed on a rotator overnight at 4 °C. Following centrifugation (27 000 *g* for 30 min at 4 °C) supernatant was mixed for 30 min with 0.5 ml DEAE-Sephacel (*Pharmacia*) equilibrated in the same denaturant [9], then centrifuged as before. The DEAE-Sephacel pellet was discarded and the supernatant diluted at least 30-fold with MOPS-metals buffer, kept for 1 h at room temperature, then centrifuged as before.

Dextran affinity chromatography [1]: either cell lysate containing *de novo* soluble product or product solubilized from pelleted cell debris by denaturant and refolded (as above), was passed down a column (10 x 65 mm bed) of Sephadex G-75 (*Pharmacia*, dry particle size 40 – 120 μm). A new column of affinity matrix was used for each separation to avoid cross-contamination. The sample was loaded at a low flow rate (0.05 – 0.1 ml/min) and the column washed (0.5 – 1 ml/min) to base line A_{280} with MOPS-metals buffer, then eluted (0.2 ml/min) with 1 or 10 mM methyl α-D-mannopyranoside in the same buffer and any A_{280} peak collected.

There are some problems with this method regarding the reproducibility, the quality and quantity of the recombinant protein produced. Stained SDS-PAGE (Sodium Dodecyl Sulphate – Poly Acrylamide Gel Electrophoresis) indicates that usually one protein species is isolated and Western blotting and N-terminal protein sequencing identify this as a form of Con A [5]. However, the yield of recombinant product is low and unreliable. The purity is called into question by broadening of the

monosaccharide-elution A_{280} profiles during affinity chromatography (Fig. 1), relative to the sharp profiles indicating homogeneous binding which are observed with authentic native Con A samples. Sometimes chromatography columns can become blocked by formation of some insoluble material or eluted fractions may show turbidity. Furthermore, the UV absorption spectrum of the material obtained (Fig. 2) indicates the presence of nucleic acid-derived contaminants resulting in over-estimation of the protein concentration relative to that expected from N-terminal sequencing results [5].

The isolation procedure was based on earlier methods [9] for recombinant pea lectin, and was also used by other workers on this protein [10], [11], [12]. Native pea lectin has similar properties to Con A and it would seem probable that similar problems might be encountered with this procedure as used with recombinant pea lectin. There appears to be no evidence that such possible inconsistencies have been eliminated in working with proteins related to Con A. It therefore appears that there are some further possible factors which compromise the quality and quantity of recombinant products obtainable.

The problems and inconsistencies given above are not simply the result of low expression levels of recombinant product, since changing the expression plasmids to newer higher level vectors did not overcome the difficulties of purification but rather highlighted them. The cellular location of the expressed product is periplasmic [5], although similar difficulties are also expected for product accumulating in the bacterial cytoplasm. During the isolation procedure outlined above, other molecules/macromolecules present in the bacterial cell are released and carried

through the process where they can interfere with the purification of active soluble forms of Con A and contaminate and form complexes with the preparations obtained.

Evidence for the presence of interfering materials was obtained by adding pure soluble (commercially-available) native Con A to extracts of bacterial cells not capable of expressing recombinant lectin. In these "spiking" experiments, precipitates formed and poor recovery of the added protein was obtained. On addition of competing monosaccharide (methyl α-D-mannopyranoside at a concentration of 10 mM), these precipitates re-dissolved: showing that their formation was dependent on the binding activity of Con A. The interfering materials could not be removed from Con A by dialysis, including dialysis in the presence of competing monosaccharide, consistent with macromolecular properties for some of the components.

The applicant has found that one of the interfering substances is glycogen synthesized in the bacterial cell, and that this can form a complex with active Con A which may trap other molecules such as nucleic acid derivatives and other substances. The presence of glycogen would reduce the quantity of recombinant lectin isolated and compromise its quality by contaminants. However, materials other than glycogen may also cause purification problems.

Glycogen is a storage polymer of glucose units mostly linked α-(1→4) with frequent α-(1→6) branching, thus carrying multiple terminal non-reducing α-glucose units [13] which can serve as receptors for Con A through the "chain-end mechanism" of interaction [1]. Sumner and Howell [2] first showed that native Con A could precipitate glycogen and Goldstein and So [14] demonstrated that it formed precipitates with a wide range of polysaccharides including glycogen from *Escherichia coli.* This bacterium contains at least as much glycogen (2.5 % dry

weight) as it does peptidoglycan in its cell wall [15]. The accumulation of glycogen is increased in carbon-rich media and is inversely related to growth rate [16], [17]. The molecular weight of glycogen from *E. coli* ranges from 10^6 - 10^8 [15], [17]: well above the cut-off of dialysis membranes.

The high starch content of the Jackbean seed does not seem to cause similar problems in the isolation of native Con A from its natural source. This is at first surprising, since the basic chemical structure of starch and glycogen is very similar. However, in plant cells starch grains are formed which can retain the polysaccharide in an insoluble form, whereas in the bacterial cell glycogen is soluble in the cytoplasm and would be released on cell breakage. It was confirmed by experiment that Jackbean seed meal did not release soluble starch into MOPS-metals buffer, but that the insoluble material tested positive for starch with iodine. Whereas some Con A is probably lost by adsorption to starch grains during extraction from its natural source so decreasing the quantity obtained, this insoluble material would be removed by centrifugation so that appreciable amounts of starch would not be carried through the process to affect the quality of the preparation. Therefore, the problem of glycogen forming a complex with recombinant Con A is a consequence of the use of a bacterial cell as expression host.

It is desirable to produce pure soluble homogeneous active recombinant Con A, but problems arise whenever active lectin contacts some cell components, of which glycogen is a major though not the sole interfering molecule. Although there are a number of potential methods by which the product might be separated from cell components, such as making soluble extracts of periplasm only *i.e.* without releasing cytoplasmic contents, acetic acid extraction of Con A from cell debris, extraction in

the presence of competing monosaccharide, and amylase digestion of glycogen then dialysis, these have all been unsuccessful.

Expression experiments in mutant *E. coli* strains [18] which were defective in glycogen synthesis have also been carried out. When preparations were made as described earlier [5], broadened affinity elution peaks were still obtained (Fig. 3) – though different in profile than those from glycogen-containing host cells (Fig. 1). Although the UV spectral scan was improved (Fig. 4) compared to glycogen-containing cells (Fig. 2), the purity was still not comparable to native plant Con A (Fig. 2). Besides nucleic acid derivatives, other possible contaminating cell components might include lipopolysaccharides, especially from *E. coli* B-derived strains since these are possible Con A receptors [19].

The problem could be solved in principle if the interfering substances could be solubilised and removed, whilst retaining the accumulated recombinant lectin as an insoluble form, which could later be dissolved in denaturant, refolded and recovered by affinity chromatography. The major portion of recombinant lectin was generally present in the fraction insoluble in MOPS-metals buffer as used in Min *et al.* [5], but this buffer did not successfully remove the interfering substances (Figs. 1 – 4). Therefore there is a need for solution formulations and a washing regime that successfully accomplishes the removal of interfering molecules with minimal losses of desired protein product.

The possible nature of the insoluble Con A protein contained in this fraction [5] must be considered when designing such solution formulations. During high level expression of proteins in heterologous environments, especially bacterial hosts, it is

not unusual for dense aggregates of inactive recombinant protein to form. It is generally accepted that these aggregates or "inclusion bodies" are the result of aggregation of partly-folded or misfolded protein chains [20] produced at high concentrations in an environment which is not conducive to proper folding. Most examples of such insoluble misfolded aggregates have been produced in the bacterial cytoplasm, but they are also known to form when the product is secreted to the periplasm. Periplasmic inclusion bodies may be amorphous and hence more easily dispersed than highly regular dense cytoplasmic inclusion bodies [21]. The various forms of recombinant Con A have also been directed to the periplasm, as in Min *et al.* [5]. Therefore it could not be assumed that conventional isolation procedures [22] for cytoplasmic inclusions would be adequate. Furthermore, in developing jack bean seeds Con A accumulates in protein bodies [23] in an insoluble but active form (since a saline extract can bind to dextran columns) [4], [24]. Thus, the possibility that some of the recombinant lectin in the insoluble bacterial fraction is likewise correctly folded and active but has precipitated in the confines of the periplasm cannot be excluded. Any such active form may specifically bind cell components of appropriate structure, in addition to any non-specific binding which may also be shown by inactive forms of Con A. Thus there is a need to remove interfering bacterial substances from active as well as inactive Con A.

Incorrectly folded aggregated proteins are expected to remain insoluble in the absence of very harsh conditions so that *e.g.* chaotropic denaturants or very strong alkali are avoided here. Native Con A is soluble in acidic conditions, but is inactivated and aggregates at pH > 8.5 [24], [25]. Low ionic strength ($I < 0.3$) decreases the stability and solubility of native Con A [26], and divalent cations (Mn^{2+} and Ca^{2+}) are

required to maintain its activity [24]. Therefore, a buffer about pH 9, effective at low ionic strength and compatible with a metal-chelating agent, is required.

Thus, in the first aspect the present invention provides a method of purifying a recombinant glucose binding protein expressed in a non-plant host cell comprising using a low ionic strength buffer with a pH between 8.5 and 9.5. In a preferred embodiment the pH of the buffer is 9.05 to 9.25 and $I < 0.1$.

The term "low ionic strength" as used herein means $I \leq 0.3$, and preferably $I < 0.1$.

Two buffering compounds have been found that meet the criteria above: CHES {2-(cyclohexylamino)-ethanesulphonic acid} with pK = 9.50, or borate {as $Na_2B_4O_7.10\ H_2O$, disodium tetraborate decahydrate or borax with pK = 9.24 (at 25°C, [27]). Experiments at pH 9.1 using CHES showed that this approach would work, but borate gave better indications of protein purity and was preferred over CHES. Borate is used at a pH close to its pK to maximize its buffering capacity at a low ionic strength [27]. The disodium salt of ethylene-diamine-tetra-acetic acid (EDTA.Na_2) was included to chelate metal ions [27].

Thus in one preferred embodiment of the invention the buffer comprises CHES or borate.

The solution should dissolve carbohydrates including polysaccharides (glycogen) and small cell wall fragments, nucleic acids for example DNA, RNA and any fragments and derivatives, lipid materials such as lipopolysaccharides, membrane and other lipids. As the desired protein is insoluble, but the interfering materials,

especially the carbohydrates are soluble, the protein can be easily separated from the contaminants.

Borate has useful properties with regard to the above. It readily forms negatively charged soluble complexes with neutral polysaccharides and has long been used for their separation [28]. Borate is also used (together with EDTA) as a component of buffers to dissolve DNA and RNA during their electrophoretic separation [29]. EDTA is also used to liberate lipopolysaccharide from *E. coli* [19]. The non-ionic detergent Triton-X-100 was included in the first washing step to solubilise lipid materials [27].

In addition the solution should ideally provide conditions which promote the release of any specifically-bound contaminants by any active Con A present. The solution should compete with or release molecules recognised by Con A. This property should also be easily eliminated and not carried through to the affinity chromatography steps, so that it could not then prevent the desired binding to dextran beads and recovery of active refolded lectin.

Release of any specifically-recognised substances as above might be achieved by using a monosaccharide to compete for the binding site of Con A. However, another advantageous feature of borate is its ability to inhibit precipitation of polysaccharides by Con A [30] and to specifically release carbohydrates (including glycogen) tightly bound to Con A [31], [32], and bound to closely related lectins [33], [34]. Borate can also be easily removed from precipitates by aqueous washing.

In one particularly preferred embodiment the buffer comprises borate, more preferably 20 mM Borax (Figs. 5 – 8).

The term "glucose binding protein" refers to any protein, polypeptide, or fragment thereof which specifically binds to glucose, *i.e.* has a binding affinity for glucose as might be indicated by a dissociation constant (K_d) of about 10^{-2} M and all values below this.

Such proteins may include any enzymes of glucose, glycogen or starch metabolism, receptors and lectins, in particular Con A. The present invention also encompasses various forms of these proteins, such as: pre-pro-proteins, pro-proteins, mature (fully processed) proteins, and any tetrameric, dimeric or monomeric forms howsoever modified in their quaternary structure. Mutant forms of the protein that are still active, *i.e.* retain the ability to bind to glucose are also covered by the invention. Thus, for instance proteins or polypeptides which include one or more additions, deletions, substitutions or the like are encompassed by the present invention. In addition, it may be possible to replace one amino acid with another of similar "type". For instance replacing one hydrophobic amino acid with another. Such mutants can include fragments of the protein, provided that the ability to bind glucose is retained.

Furthermore, such proteins may have one or more binding sites for specific sugars (including glucose) *i.e.* they may be multivalent, where valency refers to the number of carbohydrate binding sites per protein molecule. Any modified or mutant recombinant forms where the valency has been altered or reduced below that naturally found, but where at least one active site has been retained, are covered by this invention. For instance in the case of Con A, the valency might be reduced below 4 active sites to 3, 2, or 1 active sites per protein molecule – whether or not the number of subunits was also varied. Provided one or more active sites is retained in the assembled quaternary structure regardless of the number of subunits (protomers,

monomers) comprising that structure, then the present invention encompasses any such variable and low valency forms.

In another aspect the present invention provides a recombinant glucose binding protein substantially free of glycogen, and optionally other impurities. The protein is preferably a lectin, more preferably Concanavalin A or a mutant or variable valency or low valency form thereof.

"Substantially free" as used herein means that there is less than 5% glycogen present, more preferably less than 2.5%, 1.0% or 0.5% glycogen. Most preferably the protein is contaminated with less than 1% impurities including glycogen.

The present invention will now be described in the following Example, which refer to the figures listed below.

Figure 1. Dextran affinity chromatography elution profile for recombinant mature Con A refolded from the bacterial cell fraction insoluble in MOPS-metals buffer.

The vertical axis represents Absorbance at 280 nm (A_{280}) with full scale (100%) set to 0.2 A_{280}. The horizontal axis represents elution volume with time progressing from left to right (flow rate = 0.2 ml/min, chart speed = 0.5 mm/min).

Refolded protein (obtained from100 ml bacterial culture) in MOPS-metals buffer (45 ml) was pumped slowly (0.1 ml/min) down a column of Sephadex G-75 as in Min *et al.* [5]. The column was washed to baseline A_{280} at 0.5 ml/min and then eluted as shown: firstly with 10 mM, and then with 100 mM methyl α-D-mannopyranoside at 0.2 ml/min.

A typical A_{280} profile is shown with a small broad peak eluting at the lower monosaccharide concentration and further material eluting progressively after the competing monosaccharide concentration is raised. (The soluble fraction from these cells applied directly to the affinity column produced material which could only be eluted at 100 mM methyl α-D-mannopyranoside.) Peak sizes and shapes were variable in different experiments.

Figure 2. UV spectral wavelength scan after affinity chromatography of recombinant mature Con A (refolded from the bacterial cell fraction insoluble in MOPS-metals buffer).

Peak fractions **(F4, F5)** eluted by 10 mM methyl α-D-mannopyranoside from a dextran affinity column were examined by UV spectroscopy in a Beckman DU-650 spectrophotometer using the elution buffer as a reference. A commercial sample (*Sigma* Type IV) of native Con A purified from jack beans was included as a standard **(ST)** for comparison, showing a typical protein absorption spectrum with maximal absorbance around 280 nm. Peak fractions **(F4, F5)** show a relatively flat spectrum with no such maximum. The ratio **(A_{280}/A_{260})** for each scan was **F4**, 0.92; **F5**, 0.96 and **ST**, 1.72. This indicates entrapment of nucleic acid-derived materials in these peak fractions which had been eluted from the dextran column by a monosaccharide highly specific for the binding site of Con A.

Figure 3. Dextran affinity chromatography elution profile for recombinant mature Con A refolded from the fraction insoluble in MOPS-metals buffer obtained from glycogen-deficient mutant bacterial cells.

The vertical axis represents Absorbance at 280 nm (A_{280}) with full scale (100%) set to 0.5 A_{280}. The horizontal axis represents elution volume with time progressing from left to right (flow rate = 0.2 ml/min, chart speed = 0.5 mm/min).

Refolded protein (obtained from 250 ml bacterial culture) in MOPS-metals buffer (90 ml) was pumped slowly (0.1 ml/min) down a column of Sephadex G-75 as in Min *et al.* [5]. The column was washed to baseline A_{280} at 0.5 ml/min and then eluted with 10 mM methyl α-D-mannopyranoside at 0.2 ml/min.

A typical A_{280} elution profile is shown with a broadened noisy peak. (The soluble fraction applied directly to the affinity column produced small peaks eluting at 10 mM methyl α-D-mannopyranoside if glycogen-deficient mutant cells were grown at reduced temperature (28°C instead of 37°C).) Peak sizes and shapes were variable in different experiments.

Figure 4. UV spectral wavelength scan after affinity chromatography of recombinant mature Con A (refolded from the fraction insoluble in MOPS-metals buffer) obtained from glycogen-deficient mutant bacterial cells.

A peak fraction eluted by 10 mM methyl α-D-mannopyranoside from dextran affinity chromatography was examined by UV spectroscopy in a Beckman DU-650 spectrophotometer using the elution buffer as a reference. This scan is now closer to

pure Con A than the peak fractions shown in Fig. 2, but still shows elevated absorbances at wavelengths below 280 nm compared to pure Con A. The ratio (A_{280}/A_{260}) obtained was 1.06, well below the value for pure Con A (Fig. 2). Elimination of glycogen synthesis from the host cells expressing recombinant Con A appears to have improved the product, but not solved all the problems of contaminating materials.

Figure 5. Dextran affinity chromatography elution profile for recombinant mature Con A refolded after using the Borate Wash Method.

The vertical axis represents Absorbance at 280 nm (A_{280}) with full scale (100%) set to 0.2 A_{280}. The horizontal axis represents elution volume with time progressing from left to right (flow rate = 0.2 ml/min, chart speed = 0.5 mm/min).

Insoluble material, obtained from 100 ml bacterial culture (not a glycogen-deficient mutant) after processing by the Borate Wash Method, was refolded in MOPS-refolding buffer (45 ml) and then pumped slowly (0.1 ml/min) down a column of Sephadex G-75. The column was washed to baseline A_{280} at 0.5 ml/min and then eluted with 10 mM methyl α-D-mannopyranoside at 0.2 ml/min.

A typical A_{280} profile is shown and the sharp symmetrical peak indicates that a single homogeneous component has been specifically eluted by the competing monosaccharide. No further A_{280}-detectable material was eluted if the concentration of methyl α-D-mannopyranoside was raised to 100 mM. Sharp symmetrical peaks were consistently obtained in repeated experiments, with the shape invariant and the size (area) dependent on the amount of recombinant Con A derivative being purified.

Figure 6. UV spectral wavelength scan following affinity chromatography of recombinant mature Con A refolded after using the Borate Wash Method.

A typical peak fraction eluted by 10 mM methyl α-D-mannopyranoside from dextran affinity chromatography examined by UV spectroscopy in a Beckman DU-650 spectrophotometer using the elution buffer as a reference. The shape of the absorption spectrum, and the ratio (A_{280}/A_{260}) of 1.68, now compare very well with pure plant-derived Con A (Fig. 2). UV spectra with these characteristics were consistently obtained in repeated experiments employing the Borate Wash Method and indicate that it removes substances previously found to interfere with purification of recombinant Con A variants.

Figure 7. Stained SDS-PAGE showing repeated purification experiments on recombinant mature Con A using the Borate Wash Method.

Sodium Dodecyl Sulphate – Poly Acrylamide Gel Electrophoresis (SDS-PAGE) was used to separate proteins according to subunit molecular weight [35] on a 15% gel which was then stained with Coomassie Blue to show all proteins present. The first five lanes (starting from the left-hand side) each show individual purification experiments using the Borate Wash Method followed by dextran affinity chromatography. Only one band is seen in every case indicating that protein homogeneity has been achieved. The gel has been deliberately overloaded to demonstrate the absence of minor protein bands due to contamination or degradation

of the recombinant Con A. The last three lanes show total proteins (extracted by boiling in SDS-sample buffer) from bacterial cells expressing recombinant mature Con A. The degree of protein purification and consistency of the results is clearly evident.

Figure 8. MALDI-TOF mass spectrogram following affinity chromatography of recombinant mature Con A refolded after using the Borate Wash Method.

Matrix Assisted Laser Desorption Ionisation – Time of Flight (MALDI-TOF) mass spectrometry was used to analyse a typical peak fraction eluted by 10 mM methyl α-D-mannopyranoside from dextran affinity chromatography after using the Borate Wash Method. A Kratos Kompact MALDI 3 instrument was used in linear, positive ion mode. The sample as eluted (protein concentration = 0.12 mg/ml) was applied directly to the grid with no further treatment. Sinapinic acid was used as the matrix and data were acquired from 50 shots with a N_2-laser. Bovine Serum Albumin was used as a calibration protein (not shown).

A single species is seen with the expected mass for this version of recombinant mature Con A. It is clearly evident that there are no other molecules detected across the Mass/Charge range of 10,000 – 50,000 analysed here.

Example

Borate Wash Method for purifying recombinant Con A variants from *E. coli.*

Borate wash buffer (1 litre).

20 mM borax (7.63 g $Na_2B_4O_7.10H_2O$, disodium tetraborate decahydrate)

5 mM EDTA (1.86 g $Na_2EDTA.2H_2O$, ethylene-diamine-tetra-acetic acid, disodium salt, dihydrate)

Made up to 1 litre with water, the pH was 9.1-9.2 and did not need to be adjusted.

MOPS-metals buffer (1 litre).

1 M NaCl (58.44 g sodium chloride)

50 mM MOPS (10.5 g 3-(*N*-morpholino)propanesulphonic acid)

Made up to 1 litre with water and pH adjusted to 7.0 at room temperature using 5 M NaOH. Calcium and manganese were added (1 ml each of 1 M $CaCl_2$ and 1 M $MnCl_2$) giving a final concentration for each metal of 1 mM. (To prevent their precipitation the Ca^{2+} or Mn^{2+} must not be added before the pH is adjusted.)

Sodium azide (0.1 g NaN_3) was added as a preservative.

Refolding buffer.

The same MOPS-metals buffer made up as above was used, except that the final concentrations of calcium and manganese ions were 10 mM.

Bacterial cell collection.

The cells containing expressed recombinant Con A were harvested by centrifugation. Liquid culture (*e.g.* 500 ml) was spun at 6,400 X *g* for 15 min at 4°C. The supernatant was poured off and discarded. The pelleted cells were stored in a -80°C freezer until needed.

Extraction of cells.

N.B. During this procedure, the correct volume ratios were maintained: for the frozen cell pellet obtained from each 250 ml aliquot of bacterial culture, a wash volume of 20 ml was used in steps 1 to 8 (below), and then 3 ml guanidine-HCl were used to solubilise the pelleted material at step 10 (below). This 3 ml aliquot of extract was then refolded by 30-fold dilution in refolding buffer (90 ml).

1. The frozen cell pellet from 500 ml of liquid culture was thawed in 40 ml borate wash buffer containing 1% (v/v) Triton-X-100.

2. The cell pellet was re-suspended as much as possible using a glass rod and divided equally between two centrifuge tubes (*e.g.* clear Oakridge type, 50 ml size).

3. Cells were broken open and DNA was sheared by 3 X 1 min ultra-sonication with tubes embedded in hard-packed ice. A sonicator (*e.g.* MSE Soniprep 150) set at maximum power was used with the probe ~ 5mm from the bottom of the tube. The suspension was allowed to cool for ~ 1 min between sonication periods.

4. The cell lysate was incubated at room temperature for 10 – 60 min with gentle mixing on an oscillating platform.

5. The cell debris was centrifuged at 48,400 X *g* for 20 min at 4°C.

6. The supernatant was poured off carefully and the washed pellet was kept. The tube was drained by standing upside down on a piece of tissue paper.

N.B. The pellet was made up of two parts – a lower harder more opaque part and an upper less dense translucent part. The upper part contained most of the Con A (about 85%) and, particularly after washing with water (step 8 below), could become very light and was easily resuspended. Great care had to be taken when pouring off the

supernatant in case any of the pellet was lost. After the second (step 7) and third washes (step 8), the tube was not stood upside down on tissue as some of the pellet might have been lost.

7. Each pellet (now equivalent to material from 250 ml bacterial culture) was resuspended in 20 ml borate wash buffer <u>without</u> Triton-X-100 using 1 X 1 min sonication. It was then mixed at room temperature for 10 – 30 min, then centrifuged at 48,400 X *g* for 20 min at 4°C, and drained, but not allowed to stand upside down.

8. The pellet was resuspended in 20 ml water using 1 X 1 min sonication and immediately centrifuged at 48,400 X *g* for 20 min at 4°C. The supernatant was carefully discarded without disturbing the pellet.

9. The drained washed pellet was stored in a -80°C freezer until needed.

10. The frozen pellet was thawed and a thin glass rod was used to break it up as finely as possible to give a creamy consistency. Each pellet was quickly dissolved in denaturant (3 ml of 8 M guanidine-hydrochloride made up in refolding buffer), as follows. The guanidine-hydrochloride solution was added carefully as a layer on top of the creamed pellet and then a plastic pasteur pipette was used to mix the layers rapidly by vigorously sucking up and expelling the liquid.

11. The guanidine-hydrochloride extract was mixed at room temperature for up to 1 h to ensure complete dissolution of material from the pellet. Sonication (1 X 1 min at room temperature) was sometimes carried out if undissolved material remained visible. At this stage, the two dissolved pellets were combined giving 6 ml extract (containing material processed from 500 ml bacterial culture).

12. The guanidine extract could be used immediately or frozen at -80°C for storage.

Refolding.

1. The (thawed) guanidine extract was then refolded by rapidly diluting 30-fold into ice-cold refolding buffer: the 6 ml of guanidine extract was poured into 180 ml of refolding buffer while swirling to mix.

2. The refolding mixture was kept on ice for 1 h.

3. The flask was then allowed to stand at room temperature for 1 h to complete the formation of any precipitate.

4. The refolded extract was centrifuged at 30,100 X *g* for 20 min at 4°C, and the clear supernatant was then loaded onto a dextran affinity column, usually overnight.

Dextran affinity chromatography.

1. A glass column (*e.g. Pharmacia* type C10, for a 10 x 65 mm bed) was packed with ~ 5 ml settled bed of Sephadex G-75 (*Pharmacia*, dry particle size 40 – 120 µm) equilibrated in MOPS-metals buffer (0.5 ml/min). A new bed of affinity matrix was packed for every purification.

2. Refolded extract (*e.g.* 180 ml) was pumped down the column overnight at 0.1 – 0.2 ml/min.

3. The column was washed (0.4 – 0.5 ml/min) with at least 10-bed volumes MOPS-metals buffer (50 ml) with A_{280} monitoring showing a flat baseline throughout.

4. The recombinant Con A variant was then eluted (0.2 ml/min) by 10 mM methyl α-D-mannopyranoside in MOPS-metals buffer. Fractions (10 – 15 X 1 ml) were collected while continuously monitoring A_{280}, and peak fractions were pooled.

5. Purified protein was quantified by UV-spectrophotometry (A_{280}). An absorption spectrum scan (250 – 300 nm) and other quality checks were performed.

Results

The problems of poor quality product, inconsistent yields and generally variable results with the purification method in the prior art [5] have been outlined earlier. An instance is illustrated in Figure 1, where inhomogeneous material continues to elute from the dextran column when the competing monosaccharide concentration is raised. This is interpreted as being due to the formation of a high molecular weight complex [1] where Con A cross-links glycogen and may entrap other materials extracted from the bacterial cell. The multiple attachment sites for the dextran affinity matrix which would be available in such a complex by virtue of its Con A component would result in gradual detachment and disintegration of the complex in the presence of a sufficient concentration of competing monosaccharide. Although material can also be specifically released from dextran binding at lower eluant concentrations (Fig. 1), this clearly does not consist of protein alone (Fig. 2) and contains substances absorbing at 260 nm. This indicates that nucleic acids are associated and suggests that these complexes vary in their avidity for the dextran matrix.

The variable and generally disappointing results (Figs. 3 and 4) from glycogen-deficient expression hosts [18] were discussed earlier. Some improvement was seen compared to glycogen-containing cells, confirming the likely role of glycogen in complexing with Con A but indicating that other contaminating materials were still associating with the protein.

In contrast, the new Borate Wash Method dramatically results in elution profiles (Fig. 5) and UV spectra (Fig. 6) corresponding to those of highly purified plant-derived Con A. This method consistently produces high quality protein (Fig. 7),

which in the case of the recombinant mature form shown, does not contain any of the fragmented polypeptides present in the natural product (compare Fig. 2 of ref. [5]). The stained SDS-PAGE (Fig. 7) indicates purity with respect to protein components, but the mass spectrogram (Fig. 8) confirms that other molecules are absent.

Although the prime concern here has been to demonstrate the superiority of the new Borate Wash Method in producing a high quality product, it has also been a very important contributing factor in increasing the quantities of recombinant protein obtainable. For example, yields of highly purified recombinant mature Con A of 10 – 20 mg from 1 litre of bacterial culture have recently been routinely obtained. This represents an improvement of between one and two orders of magnitude on the yields initially obtained for recombinant forms of Con A [5].

REFERENCES

1 Goldstein, I. J. and Poretz, R. D. (1986) in The Lectins: Properties, Functions and Applications in Biology and Medicine (Liener, I. E., Sharon, N. and Goldstein, I. J., eds.), pp. 33-247, Academic Press, Orlando, London

2 Sumner, J. B. and Howell, S. F. (1936) Journal of Bacteriology **32**, 227-237

3 Agrawal, B. B. L. and Goldstein, I. J. (1967) Biochim Biophys Acta **147**, 262-71

4 Agrawal, B. B. L. and Goldstein, I. J. (1972) in Methods in Enzymology, vol. 28 (Ginsberg, V., ed.), pp. 313-318, Academic Press, New York

5 Min, W., Dunn, A. J. and Jones, D. H. (1992) EMBO J **11**, 1303-7

6 Jones, D. H. (1995) in Perspectives on protein engineering and complementary technologies 1995 (Geisow, M. J. and Epton, R., eds.), pp. 70-73, Mayflower Worldwide, Birmingham

7 Min, W. and Jones, D. H. (1992) FEBS Lett **301**, 315-8

8 Min, W. and Jones, D. H. (1994) Nat Struct Biol **1**, 502-4

9 Stubbs, M. E., Carver, J. P. and Dunn, R. J. (1986) J Biol Chem **261**, 6141-4

10 Prasthofer, T., Phillips, S. R., Suddath, F. L. and Engler, J. A. (1989) J Biol Chem **264**, 6793-6

11 van Eijsden, R. R., Hoedemaeker, F. J., Diaz, C. L., Lugtenberg, B. J., de Pater, B. S. and Kijne, J. W. (1992) Plant Mol Biol **20**, 1049-58

12 Hoedemaeker, F. J., van Eijsden, R. R., Diaz, C. L., de Pater, B. S. and Kijne, J. W. (1993) Plant Mol Biol **22**, 1039-46

13 Stryer, L. (1995) Biochemistry, W.H. Freeman, New York

14 Goldstein, I. J. and So, L. L. (1965) Arch Biochem Biophys **111**, 407-14

15 Neidhardt, F. C. (1987) in Escherichia coli and Salmonella typhimurium : cellular and molecular biology (Neidhardt, F. C., ed.), pp. 3-6, American Society for Microbiology, Washington, D.C.

16 Preiss, J. and Romeo, T. (1989) Adv Microb Physiol **30**, 183-238

17 Dawes, E. A. (1992) in Prokaryotic structure and function : a new perspective (47th Symposium of the Society for General Microbiology) (Dow, C. S., Coles, J. A. and Mohan, S., eds.), pp. 81-122, Cambridge University Press, Cambridge

18 Hengge-Aronis, R. and Fischer, D. (1992) Mol Microbiol **6**, 1877-86

19 Nikaido, H. and Vaara, M. (1987) in Escherichia coli and Salmonella typhimurium : cellular and molecular biology (Neidhardt, F. C., ed.), pp. 7-22, American Society for Microbiology, Washington, D.C.

20 Mitraki, A. and King, J. (1989) Biotechnology (N Y) **7**, 690-697

21 Bowden, G. A., Paredes, A. M. and Georgiou, G. (1991) Biotechnology (N Y) **9**, 725-30

22 Marston, F. A. O. (1986) Biochem J **240**, 1-12

23 Bowles, D. J. and Pappin, D. J. (1988) Trends Biochem Sci **13**, 60-4

24 Liener, I. E. (1976) in Concanavalin A as a Tool (Bittiger, H. and Schnebli, H. P., eds.), pp. 17-31, Wiley, London

25 Matsuura, S. and Chen, H. C. (1980) Anal Biochem **106**, 402-10

26 McKenzie, G. H., Sawyer, W. H. and Nichol, L. W. (1972) Biochim Biophys Acta **263**, 283-93

27 Dawson, R. M. C., Elliott, D. C., Elliott, W. H. and Jones, K. M. (1986) Data for Biochemical Research, Clarendon Press, Oxford

28 Northcote, D. H. (1965) in Methods in carbohydrate chemistry, vol. V. (Whistler, R. L., ed.), pp. 49-53, Academic Press, New York

29 Sambrook, J., Maniatis, T. and Fritsch, E. F. (1989) Molecular cloning : a laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

30 Svensson, S., Hammarstrom, S. G. and Kabat, E. A. (1970) Immunochemistry **7**, 413-22

31 Kennedy, J. F. and Rosevear, A. (1973) J Chem Soc [Perkin 1] **19**, 2041-6

32 Lloyd, K. O. (1976) in Concanavalin A as a Tool (Bittiger, H. and Schnebli, H. P., eds.), pp. 323-331, Wiley, London

33 Horstmann, C., Rudolph, A. and Schmidt, P. (1978) Biochem. Physiol. Pflanzen **173**, 311-321

34 West, I. and Goldring, O. (1996) Methods Mol Biol **59**, 177-85

35 Laemmli, U. K. (1970) Nature **227**, 680-5

CLAIMS

1. A method of purifying a recombinant glucose binding protein expressed in non-plant host cells comprising treating a lysate of said cells with a buffer in which glycogen is soluble, but in which said protein is insoluble.

2. A method as claimed in claim 1 wherein other impurities are also soluble in said buffer.

3. A method as claimed in claim 1 or claim 2 wherein said buffer is a low ionic strength buffer ($I < 0.3$) with a pH between 8.5 and 9.5.

4. A method as claimed in claim 3 wherein said buffer further comprises a metal chelating agent.

5. A method as claimed in claim 4 wherein said metal chelating agent is EDTA.

6. A method as claimed in any one of claims 1 to 4 wherein said buffer further comprises a non-ionic detergent.

7. A method as claimed in claim 6 wherein said non-ionic detergent is Triton-X-100.

8. A method as claimed in any one of claims 1 to 7 wherein said buffer comprises 2-(cyclohexylamino)-ethanesulphonic acid.

9. A method as claimed in any one of claims 1 to 7 wherein said buffer comprises borate.

10. A method as claimed in claim 9 wherein said buffer is 20 mM. Borax ($Na_2B_4O_7.10H_2O$.)

11. A method as claimed in any one of claims 1 to 10 wherein said pH is between 9.05-9.25.

12. A method as claimed in any one of claims 1 to 11 wherein $I < 0.1$.

13. A method as claimed in any one of claims 1 to 12 wherein said non-plant host is a bacterium.

14. A method as claimed in claim 13 wherein said bacterium is *Escherichia coli*.

15. A method as claimed in any one of claims 1 to 14 wherein said glucose binding protein is a glucose binding lectin.

16. A method as claimed in claim 15 wherein said lectin is Concanavalin A.

17. A protein isolated by a method as defined in any one of claims 1 – 16.

18. The use of a buffer in which glycogen is soluble, but in which a glucose binding protein is insoluble in the purification of a recombinant glucose binding protein expressed by a non-plant host cell.

19. The use as claimed in claim 18 modified by any of the features as claimed in any one of claims 2 – 16.

20. A recombinant glucose binding protein which is substantially free of glycogen, and optionally other impurities.

21. A protein as claimed in claim 20, wherein said protein is a lectin.

22. A protein as claimed in claim 21, wherein said lectin is Concanavalin A, or a mutant, or a variable valency or low valency form thereof.

Figure 1.

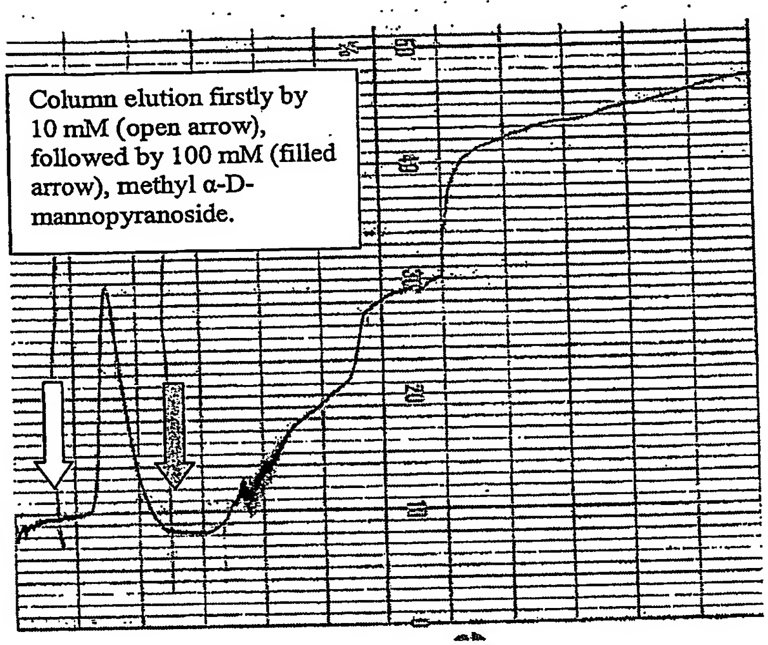


Figure 2.

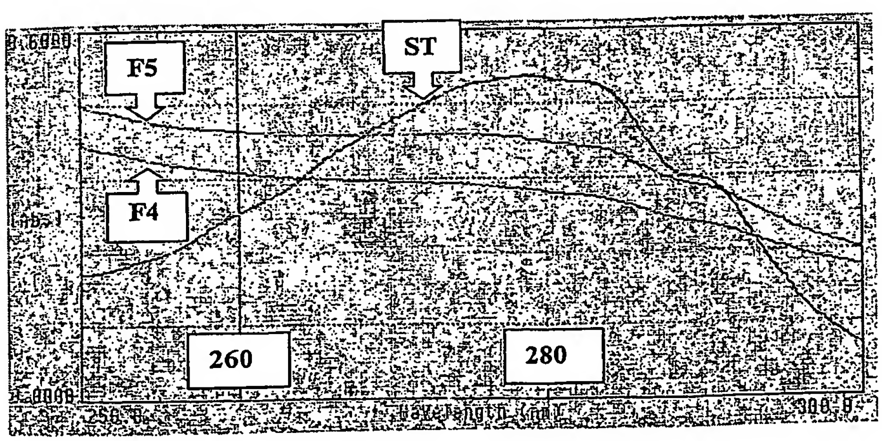


Figure 3.

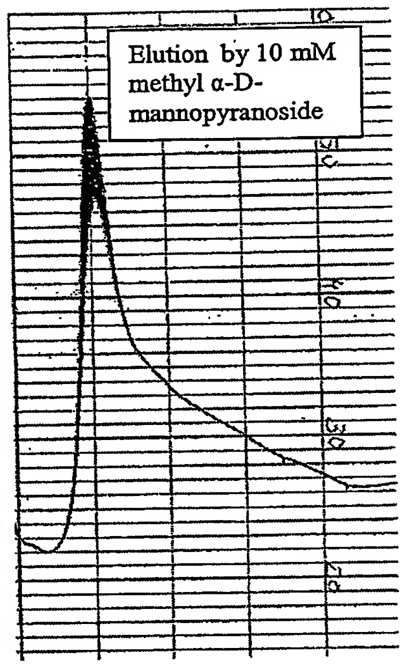


Figure 4.

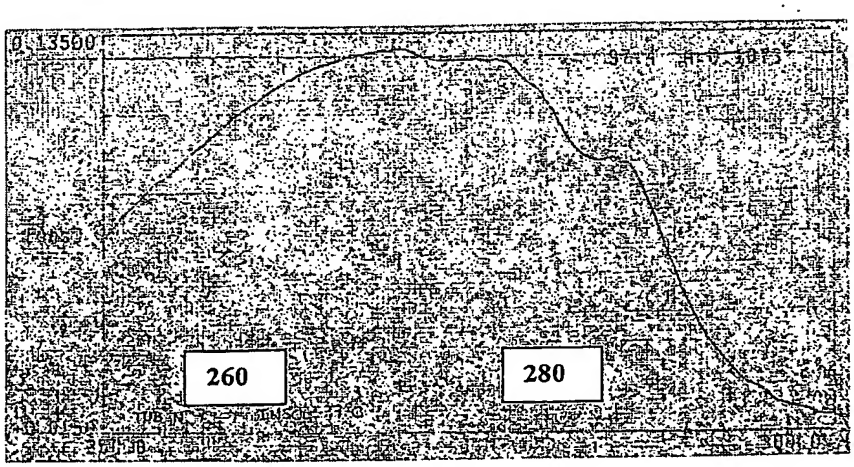


Figure 5

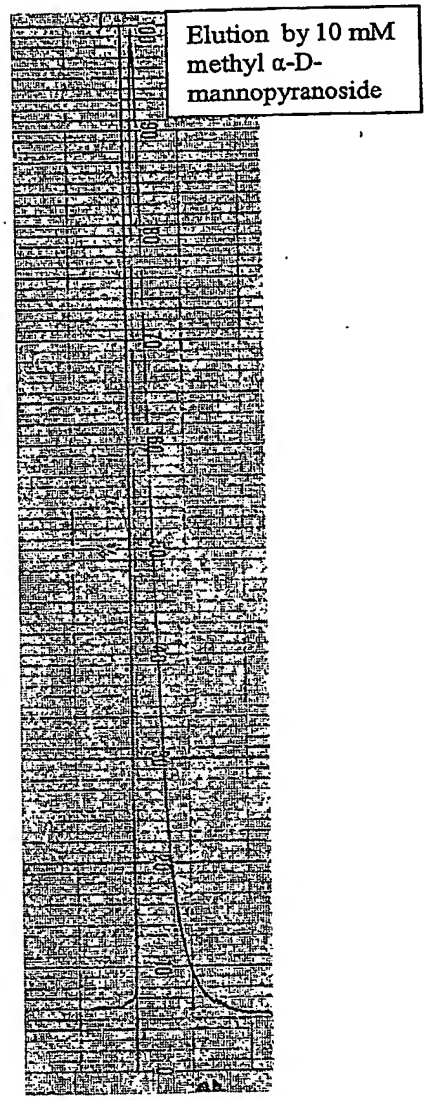


Figure 6.

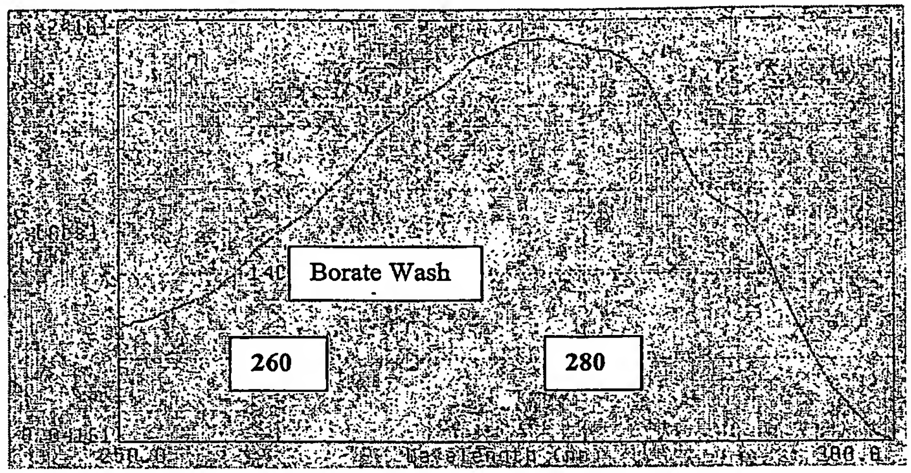


Figure 7.

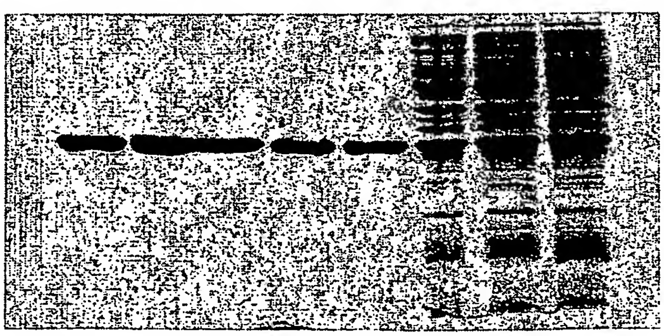
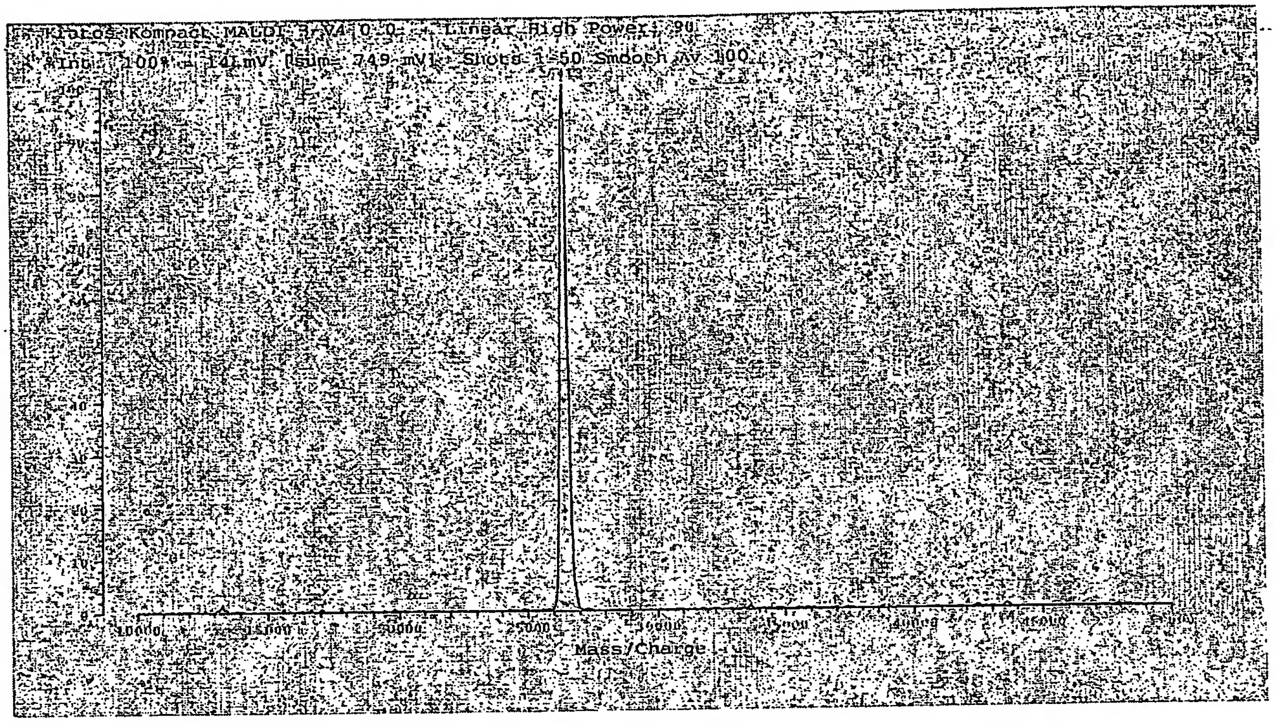


Figure-8.



PCT/GB2004/004907